

# VU Research Portal

## Introducing the human factor in predictive modelling: a work in progress

Verhagen, J.W.H.P.; Nuninger, L.; Tourneux, F.-P.; Bertoncello, F.; Jeneson, K.

### ***published in***

Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012  
2013

### ***document version***

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### ***citation for published version (APA)***

Verhagen, J. W. H. P., Nuninger, L., Tourneux, F.-P., Bertoncello, F., & Jeneson, K. (2013). Introducing the human factor in predictive modelling: a work in progress. In G. Earl, T. Sly, A. Chrystanthi, P. Murrieta-Flores, C. Papadopoulos, I. Romanowska, & D. Wheatley (Eds.), *Archaeology in the Digital Era. Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012* (pp. 379-388). Amsterdam University Press.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Introducing the Human Factor in Predictive Modelling: a Work in Progress

**Philip Verhagen**

*VU University Amsterdam, The Netherlands*

**Laure Nuninger**

*CNRS, UMR6249, Besançon, France*

**François-Pierre Tourneux**

*Université de Franche-Comté, Besançon, France*

**Frédérique Bertoncello**

*CNRS, UMR 7264, Nice, France*

**Karen Jeneson**

*Thermenmuseum, Heerlen, The Netherlands*

## **Abstract:**

*In this paper we present the results of a study aiming at integrating socio-cultural factors into predictive modelling. So far, predictive modelling has largely neglected the social and cultural dimensions of past landscapes. To maintain its value for archaeological research, therefore, it needs new methodologies, concepts and theories. For this study, we have departed from the methodology developed in the 1990s during the Archaeomedes Project. In this project, cross-regional comparisons of settlement location factors were made by analyzing the environmental context of Roman settlements in the French Rhône Valley. For the current research, we expanded the set of variables with 'socio-cultural' factors, in particular accessibility, visibility, and the effect of previous occupation, and created predictive models from this. In this way, we have developed a protocol for predictive modelling using both environmental and socio-cultural factors that can easily be implemented for different regions and time periods.*

## **Keywords:**

*Predictive Modelling, Socio-Cultural Factors, Regional Comparison, Diachronic Comparison, Roman Period*

## **1. Introduction**

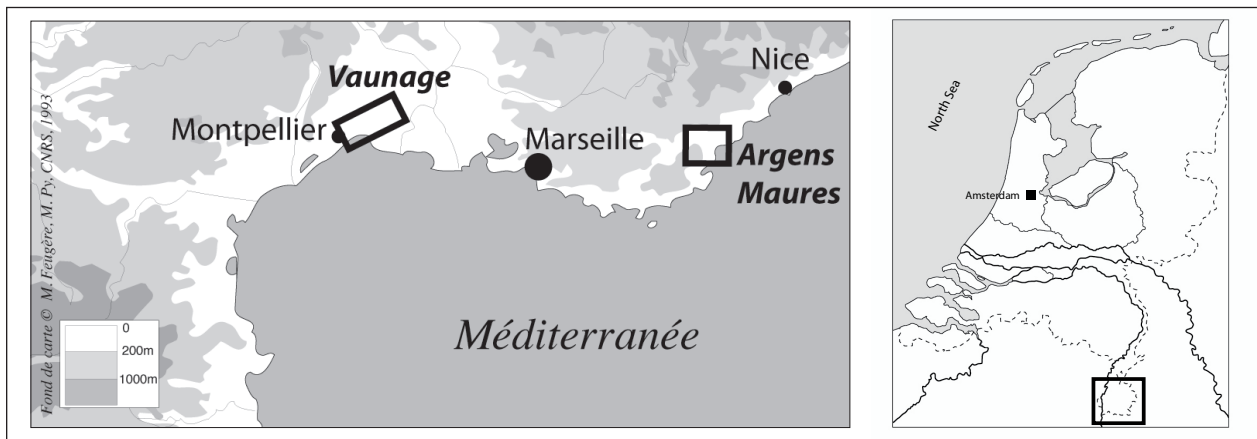
Archaeological predictive modelling has a long history of application, especially in cultural resources management (see Judge and Sebastian 1988; Verhagen 2007; Kamermans et al. 2009). Despite its popularity for archaeological heritage management, it has also been the subject of substantial criticism from academic researchers (van Leusen 1996; Wheatley 2004; van Leusen and Kamermans 2005; Kamermans 2007). The goals of predictive modelling in heritage management are the accurate and cost-effective prediction of the location of archaeological remains within a limited region. However, academic researchers are usually more interested in finding explanations of why

archaeological remains are concentrated in particular parts of the landscape. Predictive modelling can be used as a tool for this purpose as well, but should be used with caution. Little attention is paid to the role of socio-cultural factors in prehistoric and historical site location choice (Verhagen et al. 2010). The result is a rather static way of modelling, in which the human factor remains unexplored. Furthermore, issues of temporality have been addressed uncritically or insufficiently. To maintain its value for archaeological research, therefore, predictive modelling needs new methodologies, concepts and theories.

For this study, we have tried to address this issue by developing a protocol for predictive modelling, using both environmental and socio-cultural factors, that can easily be implemented for

---

*Corresponding author: j.w.h.p.verhagen@vu.nl*



**Figure 1.** Location of the study areas in France and the Netherlands.

different regions and time periods. The methodology used was originally developed in the 1990s during the Archaeomedes Project (Favory et al. 1999; van der Leeuw 1998; van der Leeuw et al. 2003; Nuninger et al. 2008). In this project, the surroundings of Roman settlements in 8 study regions in the French Rhône Valley were analyzed in order to make cross-regional comparisons of potential settlement location factors, like slope, aspect, solar radiation and soil type. However, at the time this method was not used for predictive modelling purposes, and did not include an analysis of non-environmental factors. For the current research, we have adapted this approach and applied it in three different study regions, the Vaunage (Languedoc, France), Argens-Maures (Provence, France) and Zuid-Limburg (the Netherlands; Fig. 1).

When we are speaking of ‘socio-cultural’ variables, we can distinguish two different classes. On the one hand, we have variables that can be described as measurable attributes of the archaeological site sample itself; they are not related to any environmental factor. These include properties like site location, size, functional type or duration of occupation. They can be considered as expressions of forms of social behaviour, even when the interpretation of the specific behaviour involved may be subject to discussion. In themselves, these variables are not extremely difficult to obtain, but the problems of analysing and interpreting archaeological site databases are manifold and must be addressed before these properties can actually be used for predictive modelling. The second class of socio-cultural variables concerns features of the

landscape itself that can be interpreted as having socio-cultural significance, for example visibility, topographical prominence, or accessibility. These could be described as socio-cultural landscape variables, and are not necessarily excluded from ‘traditional’ predictive modelling. It can, in fact, be argued that all environmental variables have a cultural component, even though for most variables used in traditional predictive modelling this cultural aspect is usually related to subsistence economy, and not to e.g. ritual practices.

When thinking about ways to include socio-cultural factors in predictive modelling, two main approaches can be followed, that can be described as theory-driven (deductive) and data-driven (inductive) modelling (see Verhagen 2007). Both approaches have been applied to predictive modelling ever since the technique was developed in the 1970s, but data-driven modelling has long been dominant (see also Judge and Sebastian 1988; van Leusen and Kamermans 2005). Data-driven models take a set of input variables, usually representing natural landscape features, and use statistical methods to find correlations between archaeological site location and the environment. These correlations are then extrapolated to areas where no archaeological sites have been found. Theory-driven models on the other hand, depart from hypotheses about site location preferences, and combine the variables involved, usually by some form of multi-criteria analysis, into a predictive model. Theory-driven models have clear advantages over data-driven models: they involve the perspective of human decision making in



**Figure 2.** *The landscape characteristics of the study areas.*

modelling, and there is no need to develop elaborate archaeological site databases in order to run and test the models (Verhagen and Whitley 2012). There are however also disadvantages to this type of modelling, because it needs a well-defined theoretical framework for settlement location choice. Whitley et al. (2010), for example, developed an elaborate theory-driven predictive model for settlement location choice in the coastal areas of Georgia (USA), based on detailed historical and ecological information on subsistence strategies of the 16th century native American population in the area. In many cases, these frameworks are not available or insufficiently developed, or the necessary data may not be available. So, in practice, many predictive models will continue to be made using a data-driven approach, trying to get the most out of the available environmental and archaeological datasets. In those cases, the inclusion of socio-cultural factors can still be an option, but very few studies have actually tried to do this (but see Ridges 2006). The current paper takes a somewhat intermediate position, and uses theoretically informed data-driven modelling to better understand the role of socio-cultural variables for Roman rural settlement patterns.

## **2. Study Areas**

The study areas were chosen because they all have good quality archaeological data of rural settlements in the Roman period. This allows us to make cross-regional comparisons and test the quality of the predictive models made. Two of those areas, the Vaunage and Argens-Maures, are regions that were already analyzed during the Archaeomedes Project. The data for Zuid-Limburg are part of a larger dataset stretching into adjacent areas of Belgium and Germany that was analyzed by Karen Jeneson for her PhD-research (Jeneson in prep.).

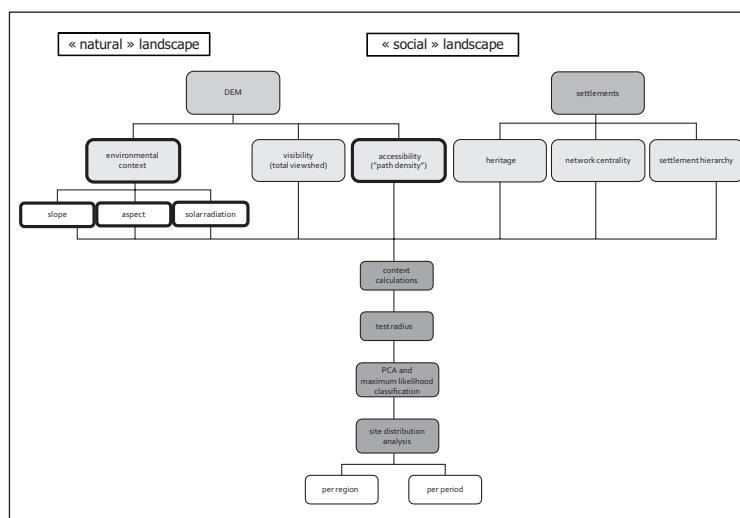
The study regions show considerable differences in landscape (Fig. 2). The Vaunage region is a basin surrounded by low hills in the Languedoc west of the town of Nîmes, with a total area of 204 km<sup>2</sup>. The Argens-Maures region consists of a mountain range (Massif des Maures) on the Mediterranean coast near Fréjus, with the Argens river valley running on its north side from east to west. The total area is 986 km<sup>2</sup>. Zuid-Limburg is a hilly region of 914 km<sup>2</sup>, with the river

Meuse running from south to north on its western margin. The major towns in it are Maastricht and Heerlen. The original research area stretches across the Dutch border into Belgium and Germany, but it was decided to restrict the study region to the Dutch part, since it was impossible to obtain environmental datasets at the same level of resolution for all three countries involved.

The archaeological data collected in the French study regions using systematic survey during the Archaeomedes Project have gone through a process of careful inspection and classification into functional categories on the basis of the available finds information (predominantly building material and site size). It allows for a chronological resolution of 100 years for most settlements. The dataset includes sites dating from 800 BC until 800 AD, so it covers a much longer time span than the Roman period. For Zuid-Limburg, the available information on site dating and function is much less detailed, no systematic survey was done, and in many cases the chronological resolution cannot be made more specific than 'Roman period' (12BC – 450AD). Nevertheless, within the Netherlands it is probably the most detailed and complete dataset available. We did not analyze geographical research biases for the three study regions, basically because of a lack of reliable evidence concerning this. The available evidence however does not indicate serious biases in data collection (Nuninger 2002), so we expect the data to be representative for all landscape contexts, except locally as in the Argens alluvial plain (Bertoncello 2011) - but this judgement is based on expert opinion.

### 3. Modelling Approach

Our main goal was to establish whether including socio-cultural factors actually makes a difference for the interpretation of site location patterns and predictive model quality. The modelling was set up using a restricted set of variables based on the available digital elevation models and archaeological settlement data. While soil type is known to have been a factor for Roman rural settlement location choice as well (Favory and



**Figure 3.** The modelling protocol. Outlined in bold are the factors that were actually analyzed for this paper.

van der Leeuw 1998), synchronizing soil mapping systems between different areas of France and the Netherlands would have been too complicated. The 50 x 50 m DEMs used for the French regions (originally obtained from the IGN) only contained integer values for elevation, and were therefore re-interpolated to floating point to get rid of the artificial terraces. For Zuid-Limburg, the impact of modern urbanisation and mining is much more evident than for the more rural French study regions. The modern-day disturbances contained in the available 5 x 5 m LiDAR-based DEM<sup>2</sup> proved to be too disruptive for the analysis, so a new DEM was created by digitizing and interpolating the 2.5 m interval contours from 1920s topographical maps scale 1:25,000<sup>3</sup>. These contours represent the situation just before large scale mining and urbanisation in the area started.

From the digital elevation models, three basic 'environmental factors' were derived: slope, aspect and solar radiation. These are thought to have been important for site location primarily because of their interest for agricultural production, and were already shown to have influenced site location patterns in the Archaeomedes Project. Furthermore, the DEMs were used to extract two additional factors that have a stronger socio-cultural connotation:

2 Actueel Hoogtebestand Nederland or AHN ([www.ahn.nl](http://www.ahn.nl))

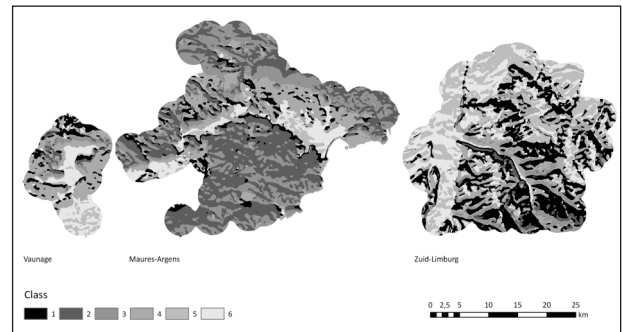
3 Chromotopografische Kaart des Rijks or Bonnebladen, the first detailed topographical maps of the Netherlands, made between the 1890s and 1930



the accessibility of the landscape (approached through the calculation of path density maps; see section 4), and visibility (through the calculation of total viewsheds). From the settlement data, three additional factors were derived that are potentially of interest to settlement location choice: the impact of previous settlement ('heritage'), the position of sites in the settlement network, and the hierarchical position of the settlements (Fig. 3). In this paper however, we will only focus on the combination of the accessibility factor with the environmental factors, since the other analyses are still in progress. It illustrates the general process of model building, and gives some preliminary answers to the question if socio-cultural factors actually improve the predictive models and offer a new perspective on site location patterns in these particular archaeological and environmental settings.

An important element of the modelling is the concept of context: instead of just analyzing the landscape characteristics at the location of an individual grid cell, a radius around each cell is used. This method was originally developed in the Archaeomedes Project, and originally only involved the analysis of radii around settlements (Favory et al. 1999; van der Leeuw 1998; van der Leeuw et al. 2003; Nuninger et al. 2008). For predictive modelling purposes however, it is also important to know the characteristics of locations where no archaeological evidence is found, so for the current study all grid cells in the study regions were taken into account. First, it involves the categorization of the variables used into discrete classes. For example, the slope map was reclassified into 5 classes (0-2%, 2-4%, 4-8%, 8-15% and > 15%) that are roughly equally distributed over the total of the three study areas. For each cell in the study areas, the proportions of these discrete classes within a certain radius are then calculated.

It should be stressed that the choice of the size of the analysis radius is not based on theoretical considerations about site catchment sizes. Instead, the appropriate radius for analysis was defined as the one that provided most statistical contrast in the context profiles. If a category was present in more than 50% of the cases (i.e. context units centred on each cell), and if the standard deviation of the proportions for this category was > 15%, then the variable was considered significant. The radius with



**Figure 4.** *Maximum Likelihood Classification of the landscape into environmental factors.*

the largest number of significant variables was then considered the most appropriate one. For each of the three study regions the most appropriate radius obtained was 250 m, corresponding to a circular region around each cell location containing 81 grid cells of 50 x 50 m.

In the first run of the model, only the environmental factors slope, aspect and solar radiation were analyzed, to provide a purely environmentally based model as a baseline. The context profiles were calculated for 16 variables for each cell in each region, a total of 819,398 grid cells. These profiles were then used to classify the landscape into broad, discrete regions sharing similar environmental characteristics. For this, first a Principal Components Analysis (PCA) was done, to identify and isolate the main statistical trends. On the basis of inspection of the scree plot of the PCA, it was decided to keep the 5 most important factors explaining almost 80% of the variation. A Maximum Likelihood Classification<sup>4</sup> was done on these 5 factors, resulting in a map with 6 different landscape classes (Fig. 4). The resulting classes can be described in general terms as follows:

Class 1: predominantly north-facing medium slopes, low solar radiation (west- and east-facing slopes represented as well).

Class 2: steep slopes, predominantly low solar radiation, but sometimes with high solar radiation as well.

<sup>4</sup> The PCA and Maximum Likelihood Classification were done in ArcGIS 9.3, using the Isocluster command to create signatures, and inspecting the resulting dendrogram to decide into how many classes the landscape should be classified.

| Region         | Period                      | $\chi^2$ | p      | max. gain | # settl. | % settl. | # contexts (cells) | % contexts (cells) |
|----------------|-----------------------------|----------|--------|-----------|----------|----------|--------------------|--------------------|
| <b>All</b>     | 800 BC - 800 AD             | 57.98    | 0.0000 | 9.7%      | 1178     | 100%     | 819277             | 100%               |
| <b>Limburg</b> | Roman (12 BC - 450 AD)      | 7.97     | 0.1578 | 4.4%      | 378      | 100%     | 352971             | 43%                |
|                | Early Roman (12 BC - 70 AD) | 5.29     | 0.3100 | 20.1%     | 26       | 7%       |                    |                    |
|                | Middle Roman (70 - 270 AD)  | 4.35     | 0.5000 | 6.9%      | 54       | 14%      |                    |                    |
| <b>Vaunage</b> | 800 BC - 800 AD             | 17.84    | 0.0030 | 8.0%      | 336      | 100%     | 81418              | 10%                |
|                | 200 - 101 BC                | 11.59    | 0.0400 | 27.7%     | 29       | 9%       |                    |                    |
|                | 100 - 1 BC                  | 7.49     | 0.1868 | 10.2%     | 75       | 22%      |                    |                    |
|                | 1 - 99 AD                   | 22.11    | 0.0005 | 14.5%     | 181      | 54%      |                    |                    |
|                | 100 - 199 AD                | 20.57    | 0.0010 | 14.9%     | 160      | 48%      |                    |                    |
| <b>Argens</b>  | 800 BC - 800 AD             | 24.47    | 0.0002 | 10.4%     | 464      | 100%     | 384888             | 47%                |
|                | 200 - 101 BC                | 6.86     | 0.2311 | 13.5%     | 76       | 16%      |                    |                    |
|                | 100 - 1 BC                  | 1.85     | 0.8697 | 3.1%      | 134      | 29%      |                    |                    |
|                | 1 - 99 AD                   | 59.43    | 0.0000 | 23.7%     | 205      | 44%      |                    |                    |
|                | 100 - 199 AD                | 62.68    | 0.0000 | 25.4%     | 189      | 41%      |                    |                    |

**Table 1.** Site distribution and predictive value for the model based on environmental factors.

Class 3: south-facing steep slopes, very high solar radiation.

Class 4: predominantly south-facing gentle slopes, high solar radiation (west- and east-facing slopes represented as well).

Class 5: flat areas or gentle slopes, low to medium solar radiation.

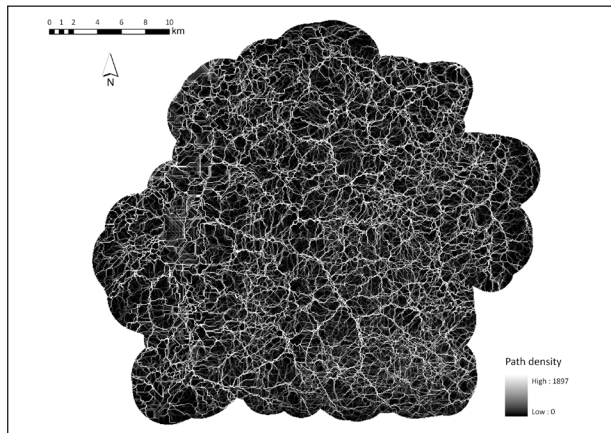
Class 6: flat areas or gentle slopes, medium to high solar radiation.

Importantly, these landscape classes are based on the total number of context profiles over all three study regions. All classes occur in all three regions, but not in equal amounts, since the landscapes characteristics are different. Since the Vaunage region is much smaller than the other two regions, its contribution to the final classification is of course less important.

The archaeological sites were then overlain on the landscape classification map, and a  $\chi^2$ -test was employed to see if any significant site location preferences could be established. Relative gains (Wansleebe and Verhart 1992; Verhagen 2007) were calculated to see how strong site location

preferences are (Table 1).

For Zuid-Limburg, the results are not statistically significant, and do not provide any evidence for a preference of settlement in a specific environmental context. For the Vaunage, the results are statistically significant and indicate an absence of preference as well. For the Argens-Maures region, a stronger patterning can be discerned. This is probably due to the more rugged relief in this region. The steeper slopes in this region will have made settlement more difficult than in the other two study areas. When looking at a more detailed chronological scale, we can see that for the French study regions the predictive power of the model is higher for the 1<sup>st</sup> century AD than for the 1<sup>st</sup> century BC, the first period of Roman settlement. It seems that during the initial agrarian colonization of the areas, environmental factors did not play an important role for choosing settlement locations. Later however, it seems that the environmental context became more important and settlements were located in more specific locations. This phenomenon was already observed in the Archaeomedes Project (Favory et al. 1999; Tourneux 2000), as well as in the Combas region north of the Vaunage by Fovet (2005). For Zuid-Limburg, this tendency seems to be reversed, with a more specific location choice in the first phase of



**Figure 5.** Path density map for Zuid-Limburg.

colonization, and an occupation of all environmental niches in the Middle Roman period. However, given the low number of accurately dated sites for this region, the reliability of this interpretation may be questioned.

#### 4. Including Accessibility

Accessibility of the landscape is a variable that can be defined in a number of ways. For this study, we used the method developed by Verhagen (in press). For a number of sample points that are distributed systematically at a distance of 250 m in each study region, accumulative cost surfaces are calculated that give the cost distance to each sample point for each grid cell. The costs used are based on Tobler's (1993) hiking equation, in which slope is determining the speed of movement by foot<sup>5</sup>. We then used the accumulative cost surfaces to calculate least cost paths to the sample points, starting from 72 locations radially distributed on the edge of a circle with a 5000 m radius from the sample point. We then added the resulting least cost paths for each sample point to obtain a cumulative cost path, or **path density** grid, for the whole area. These cumulative cost paths represent the preferred movement axes in the landscape when moving over a distance of 5000 m, given the assumption of least effort (Fig. 5). However, alternative methods could be used as well, like the total path costs developed by Llobera (2000), or the potential path fields by Mlekuž (in prep.).

<sup>5</sup>  $W = 6e^{3.5|s+0.05|}$ , where

$W$  = walking speed in km/h;  $e$  = the base of natural logarithms;  
 $s$  = slope in m/m

Due to the limitations of available computing power, the path density maps have not been calculated for the Argens-Maures region yet, and we can therefore only present the results for the Vaunage and Zuid-Limburg. The path density maps were classified into 5 discrete categories that are roughly equally distributed over both regions, and the context profiles of these categories were calculated within a radius of 250 m for each grid cell, in the same way as it was done for the environmental variables. The path density context profiles were then added as extra variables, and a new PCA and Maximum Likelihood Classification were performed on the extended set of variables. After inspection of the scree plot, the first 6 components of the PCA were retained, and the classification resulted in 5 classes.

Class 1: predominantly north-facing medium slopes, low solar radiation (west- and east-facing slopes represented as well).

Class 2: flat areas or gentle slopes, low to medium solar radiation.

Class 3: flat areas or gentle slopes, medium to high solar radiation.

Class 4: predominantly south-facing medium slopes, high solar radiation (west- and east-facing slopes represented as well).

Class 5: north-facing medium slopes, low to very low solar radiation.

The characteristics and geographical distribution of these classes are not very different from the original, environmentally based classification. The distinction between steep and medium slopes that is quite clear in the environmental classification however is less obvious in the new classification, and class 5 represents a specific category of cold, north- to northwest facing slopes. The differences in path density are not clearly reflected in the classes. This is probably due to the fact that the cumulative cost paths are relatively evenly spread over the study region. There are no distinct zones that are clearly avoided or preferred for movement. Despite the modest changes in class characteristics, a slight but significant improvement in predictive power can be observed for both regions



| Region         | Period                      | $\chi^2$ | p      | max. gain | change in p | change in max. gain |
|----------------|-----------------------------|----------|--------|-----------|-------------|---------------------|
| <b>All</b>     | 800 BC - 800 AD             | 43.27    | 0.0000 | 11.9%     | 0.0000      | 2.1%                |
| <b>Limburg</b> | Roman (12 BC - 450 AD)      | 13.05    | 0.0110 | 8.3%      | 0.1468      | 3.9%                |
|                | Early Roman (12 BC - 70 AD) | 6.93     | 0.1400 | 22.1%     | 0.1700      | 2.0%                |
|                | Middle Roman (70 - 270 AD)  | 4.06     | 0.4000 | 11.5%     | 0.1000      | 4.6%                |
| <b>Vaunage</b> | 800 BC - 800 AD             | 26.00    | 0.0000 | 11.9%     | 0.0030      | 3.9%                |
|                | 200 - 101 BC                | 2.80     | 0.5926 | 14.8%     | -0.5526     | -12.8%              |
|                | 100 - 1 BC                  | 7.59     | 0.1079 | 12.2%     | 0.0790      | 2.0%                |
|                | 1 - 99 AD                   | 22.13    | 0.0002 | 12.3%     | 0.0003      | -2.2%               |
|                | 100 - 199 AD                | 23.45    | 0.0001 | 13.7%     | 0.0009      | -1.2%               |
| <b>Argens</b>  | 800 BC - 800 AD             | 24.47    | 0.0002 | 10.4%     | 464         | 100%                |
|                | 200 - 101 BC                | 6.86     | 0.2311 | 13.5%     | 76          | 16%                 |
|                | 100 - 1 BC                  | 1.85     | 0.8697 | 3.1%      | 134         | 29%                 |
|                | 1 - 99 AD                   | 59.43    | 0.0000 | 23.7%     | 205         | 44%                 |
|                | 100 - 199 AD                | 62.68    | 0.0000 | 25.4%     | 189         | 41%                 |

**Table 2.** Site distribution and predictive value for the model including path density.

(Table 2). How to interpret this slight increase in predictive power is for the moment still open to debate, especially since it is not reflected when we look at the relative gains per period, and we have not looked at the individual contribution of path density to site location preference.

## 5. Conclusions

The results presented are preliminary, and we still need to improve on it, by including the path densities for the Argens-Maures region, and by experimenting with other approaches to accessibility that may be less demanding in terms of computing resources. Since the modelling protocol is clearly defined, it is relatively easy to include and analyze new variables. The models provide information on the relevance of variables for site location choice, and allow us to make cross-regional and diachronic comparisons of settlement pattern development. The prediction of potential settlement locations is implied, but is not the main goal of the exercise.

We want to stress that in this modelling protocol, the choice of settlement location factors for the modelling is not governed by considerations of optimal model performance, such as would be the goal of standard statistical approaches like logistic regression; the 'non-performance' of a variable is

an equally important result. Neither is this method based on an elaborate theoretical framework of Roman agricultural production, including all the possible factors that might have played a role in choosing a settlement location. Instead, we use the protocol to extract the main factors that influence settlement location over the longer term, and that have relevance for not just one study region, but for three regions in different geographical contexts. Evidently, there are limits to this approach: since the environmental side of the model is developed using elevation data, it will not work for areas that are completely flat. Similarly, certain social variables may not be relevant for other archaeological settings, or cannot even be modelled in all situations because of the limitations of the available archaeological data. For example, including the hierarchical status of settlements in the model implies that we can actually discern a hierarchy between settlements. Furthermore, the availability of good quality archaeological data sets is essential to this approach. In areas with little reliable archaeological data, a purely theory-driven approach will be more appropriate.

## Acknowledgements

The research for this paper was partly made possible through a VENI-grant awarded to Philip

Verhagen by NWO ("Introducing the human (f)actor in predictive modelling for archaeology"). Travel funds were made available to Philip Verhagen and Laure Nuninger by the Réseau Franco-Néerlandais (Van Gogh-grant) and the French Ministry of Foreign Affairs (PHC Van Gogh).

The archaeological data for Zuid-Limburg were elaborated by Karen Jeneson for her PhD-research within the framework of the research programme "Roman villa landscapes in the north: Economy, Culture, Life-Styles" (financed by NWO). The archaeological data for Argens-Maures were prepared by Frédérique Bertoncello, and for the Vaunage by François Favory, Claude Raynaud and Laure Nuninger within the frameworks of the programmes ATP-Fréjus-Argens (financed by CNRS), Archaeomedes (financed by the European Union) and Archaeodyn (financed by the Ministère de la Recherche et des Nouvelles Technologies, and later by the Agence National de la Recherche).

## References

- Bertoncello, F. 2011. "Dynamique du paysage et centuriation dans le territoire de *Forum Iulii*, Fréjus (Var, France)." In *Sistemi centuriali e opere di assetto agrario tra età romana e primo medioevo*, Atti del convegno Borgoricco (Padova), Lugo (Ravenna), 10-12 settembre 2009, edited by Pier Luigi Dall'Aglia and Guido Rosada, 75-91. Pisa: Fabrizio Serra.
- Favory, F., J.-J. Girardot, L. Nuninger, and F.-P. Tourneux. 1999. "ARCHAEOMEDES II : une étude de la dynamique de l'habitat rural en France méridionale, dans la longue durée (800 av. J.-C.-1600 ap. J.-C.)." *AGER* 9:15-35.
- Favory F., and S. van der Leeuw. 1998. "ARCHAEOMEDES, la dynamique spatio-temporelle de l'habitat antique dans la vallée du Rhône: bilan et perspectives." *Revue Archéologique de Narbonnaise* 31:257-98.
- Fovet, É. 2005. "Dynamique spatiale du peuplement et analyse des ressources agro-pastorales dans le bassin de Combas et le vallon de l'Aigalade (Gard, France), du Bronze final au bas Moyen Age." In *Temps et espaces de l'Homme en société, analyses et modèles spatiaux en archéologie : Actes des XXVe Rencontres internationales d'archéologie et d'histoire d'Antibes, 21-23 Octobre 2004*, edited by Jean-François Berger, Frédérique Bertoncello, Franck Braemer, Gourguen Davtian and Michiel Gazenbeek, 499-503. Sophia-Antipolis: Éditions APDCA.
- Jeneson, K. In prep. *Exploring the Roman rural villa world between Tongres and Cologne. A landscape archaeology approach*. Amsterdam: Vrije Universiteit Amsterdam.
- Judge, W.J., and L. Sebastian, eds. 1988. *Quantifying the Present and Predicting the Past. Theory, Method and Application of Archaeological Predictive Modeling*. Denver: U.S. Department of the Interior, Bureau of Land Management.
- Kamermans, H. 2007. "Smashing the Crystal Ball. A Critical Evaluation of the Dutch National Archaeological Predictive Model (IKAW)." *International Journal of Humanities and Arts Computing* 1:71-84.
- Kamermans, H., M. van Leusen, and P. Verhagen, eds. 2009. *Archaeological Prediction and Risk Management*. Leiden: Leiden University Press.
- Leeuw, S.E. van der, ed. 1998. *The Archaeomedes Project - Understanding the natural and anthropogenic causes of land degradation and desertification in the Mediterranean*. Luxembourg: Office for Official Publications of the European Union.
- Leeuw, S. van der, F. Favory, and J.-L. Fiches, eds. 2003. *Archéologie et systèmes socio-environnementaux: études multiscalaires sur la vallée du Rhône dans le programme ARCHAEOMEDES*. Valbonne: Presses du CNRS.
- Leusen, M. van, and H. Kamermans, eds. 2005. *Predictive Modelling for Archaeological Heritage Management: a research agenda*. Amersfoort: Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- Leusen, P. M. van. 1996. "GIS and Location Modelling in Dutch Archaeology, A Review of Current Approaches." In *New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research*, edited by H.D.G. Maschner, 177-97. Carbondale: Southern Illinois University.
- Llobera, M. 2000. "Understanding movement: a pilot model towards the sociology of movement." In *Beyond the Map. Archaeology and Spatial Technologies*, edited by G. Lock, 65-84. Amsterdam: IOS Press/Ohmsha.
- Mlekuž, D. In prep. "Exploring the topography of

- movement." In *Computational Approaches to Movement in Archaeology*, edited by S. Polla and P. Verhagen. Berlin: De Gruyter.
- Nuninger, L. 2002. *Peuplement et territoires protohistoriques du VIIIe au Ier s. av. J.-C. en Languedoc oriental (Gard-Hérault)*. Besançon: Université de Franche-Comté.
- Nuninger, L., F.-P. Tourneux, and F. Favory. 2008. "From Archaeomedes to Archaedyn." In *Layers of Perception, Proceedings of the 35th Computer Applications and Quantitative Methods in Archaeology Conference, Berlin, Germany, April 2-6, 2007*, edited by A. Posluschny, K. Lambers and I. Herzog, 278 and CD-ROM. Bonn: Dr. Rudolf Habelt GmbH.
- Ridges, M. 2006. "Regional Dynamics of Hunting and Gathering: An Australian Case Study Using Archaeological Predictive Modelling." In *GIS And Archaeological Site Location Modelling*, edited by M. W. Mehrer and K.L. Wescott, 123-43. Boca Raton: CRC Press.
- Tobler, W. 1993. *Three presentations on geographical analysis and modelling. 1) Non-isotropic modelling 2) Speculations on the geometry of geography 3) Global spatial analysis*. Santa Barbara: National Center for Geographic Information and Analysis.
- Tourneux, F.-P. 2000. *Modes de représentation des paysages*. Besançon: Université de Franche-Comté.
- Verhagen, P. 2007. *Case Studies in Archaeological Predictive Modelling*. Leiden: Leiden University Press.
- Verhagen, P. *In press*. "On the road to nowhere? Least cost paths, accessibility and the predictive modelling perspective." In *Fusion of Culture. Proceedings of the XXXVIII Conference on Computer Applications and Quantitative Methods in Archaeology, 6-9 April 2010, Granada*, edited by F. Contreras, M. Farjas and F.J. Melero. Oxford: Archaeopress.
- Verhagen, P., and T.G. Whitley. 2012. "Integrating Archaeological Theory and Predictive Modelling. A Live Report from the Scene." *Journal of Archaeological Method and Theory* 19:49-100.
- Verhagen, P., H. Kamermans, M. van Leusen, J. Deeben, D. Hallewas, and P. Zoetbrood. 2010. "First thoughts on the incorporation of cultural variables into predictive modelling." In *Beyond the artefact – Digital Interpretation of the Past - Proceedings of CAA2004 - Prato 13-17 April 2004*, edited by F. Niccolucci and S. Hermon, 307-11. Budapest: Archaeolingua.
- Wansleeben, M., and L.B.M. Verhart. 1992. "The Meuse Valley Project: GIS and site location statistics." *Analecta Praehistorica Leidensia* 25:99-108.
- Wheatley, D. 2004. "Making Space for an Archaeology of Place." *Internet Archaeology* 15. [intarch.ac.uk/journal/issue15/wheatley\\_index.html](http://intarch.ac.uk/journal/issue15/wheatley_index.html).
- Whitley, T. G., G. Moore, G. Goel, and D. Jackson. 2010. "Beyond the Marsh: Settlement Choice, Perception and Spatial Decision-Making on the Georgia Coastal Plain." In *Making History Interactive. Computer Applications and Quantitative Methods in Archaeology (CAA). Proceedings of the 37th International Conference, Williamsburg, Virginia, United States of America, March 22-26, 2009*, edited by B. Frischer, J. Webb Crawford and D. Koller, 380-90. Oxford: Archaeopress.